DeepLight: Learning Illumination for Unconstrained Mobile Mixed Reality

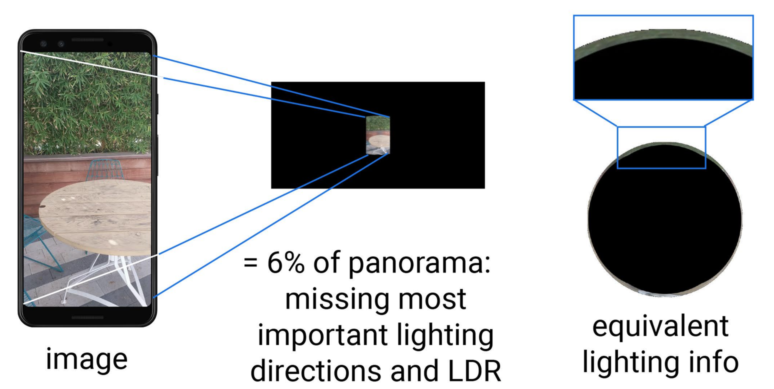
Chloe LeGendre[1, 2] Wan-Chun Ma[1] Graham Fyffe[1] John Flynn[1] Laurent Charbonnel[1] Jay Busch[1] Paul Debevec[1]
1. Google 2. USC Institute for Creative Technologies

1. Problem: How to light virtual assets for mobile Augmented Reality (AR)?

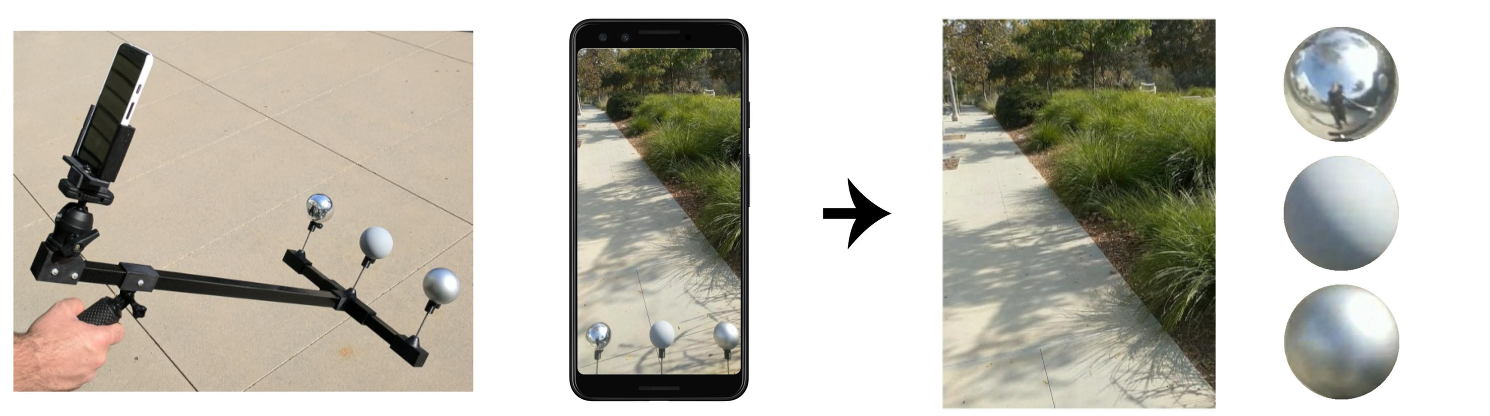
For realistic AR, the lighting used to render virtual objects must be consistent with the lighting in the real-world scene.

Real world lighting: 360° and High Dynamic Range (HDR) [1].

Mobile phone images: Low Dynamic Range (LDR) with limited FOV.



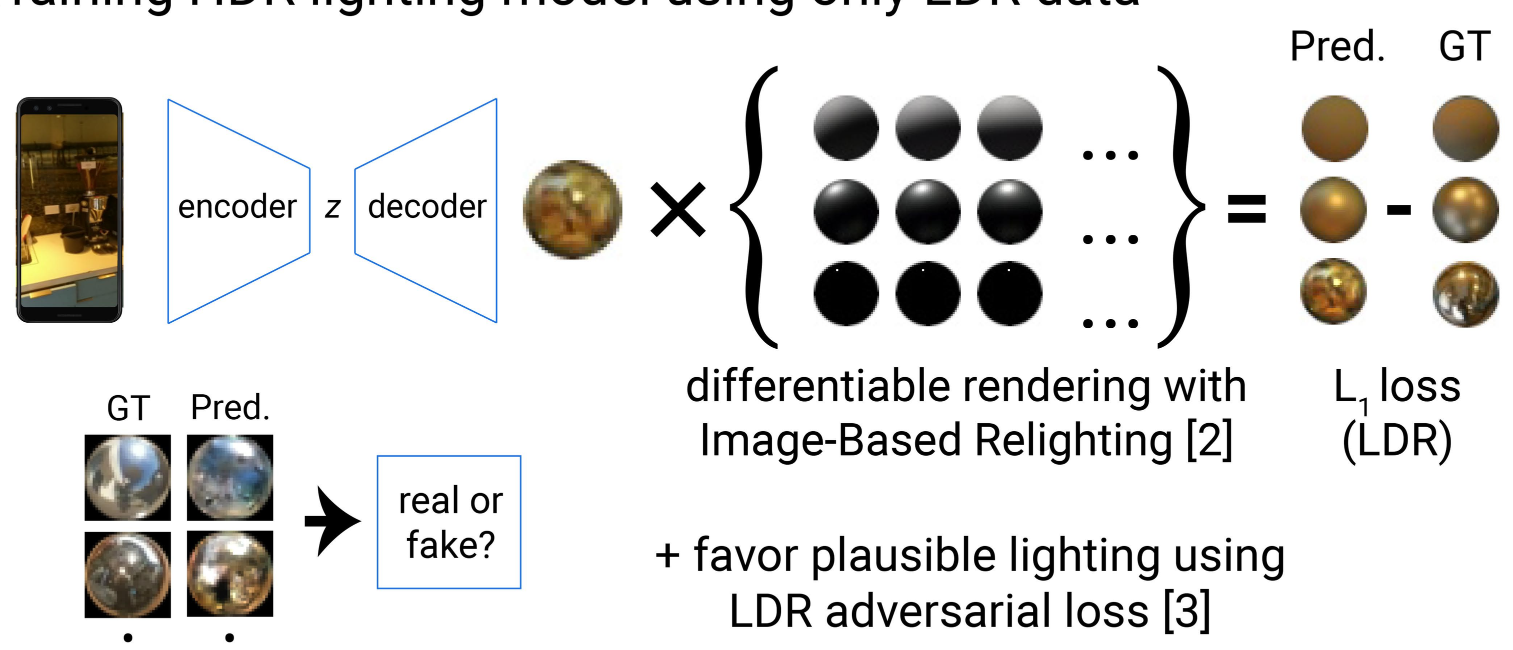
2. Approach: Learning illumination from a single image.



training data capture rig capture LDR video

training data: paired background images and ground truth LDR sphere images for 3 BRDFs

Training HDR lighting model using only LDR data



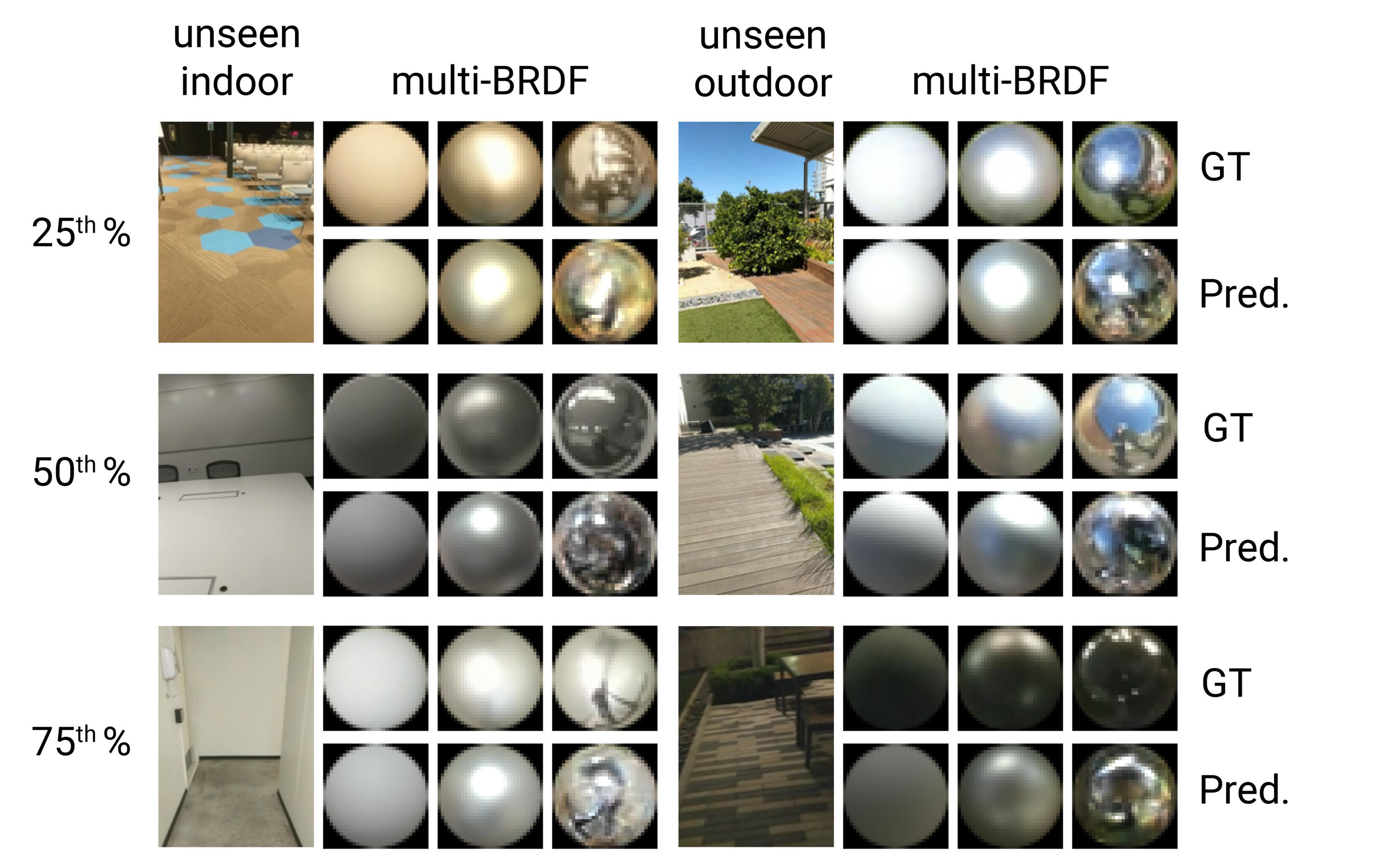
differentiable rendering with Image-Based Relighting [2]

L_1 loss (LDR)

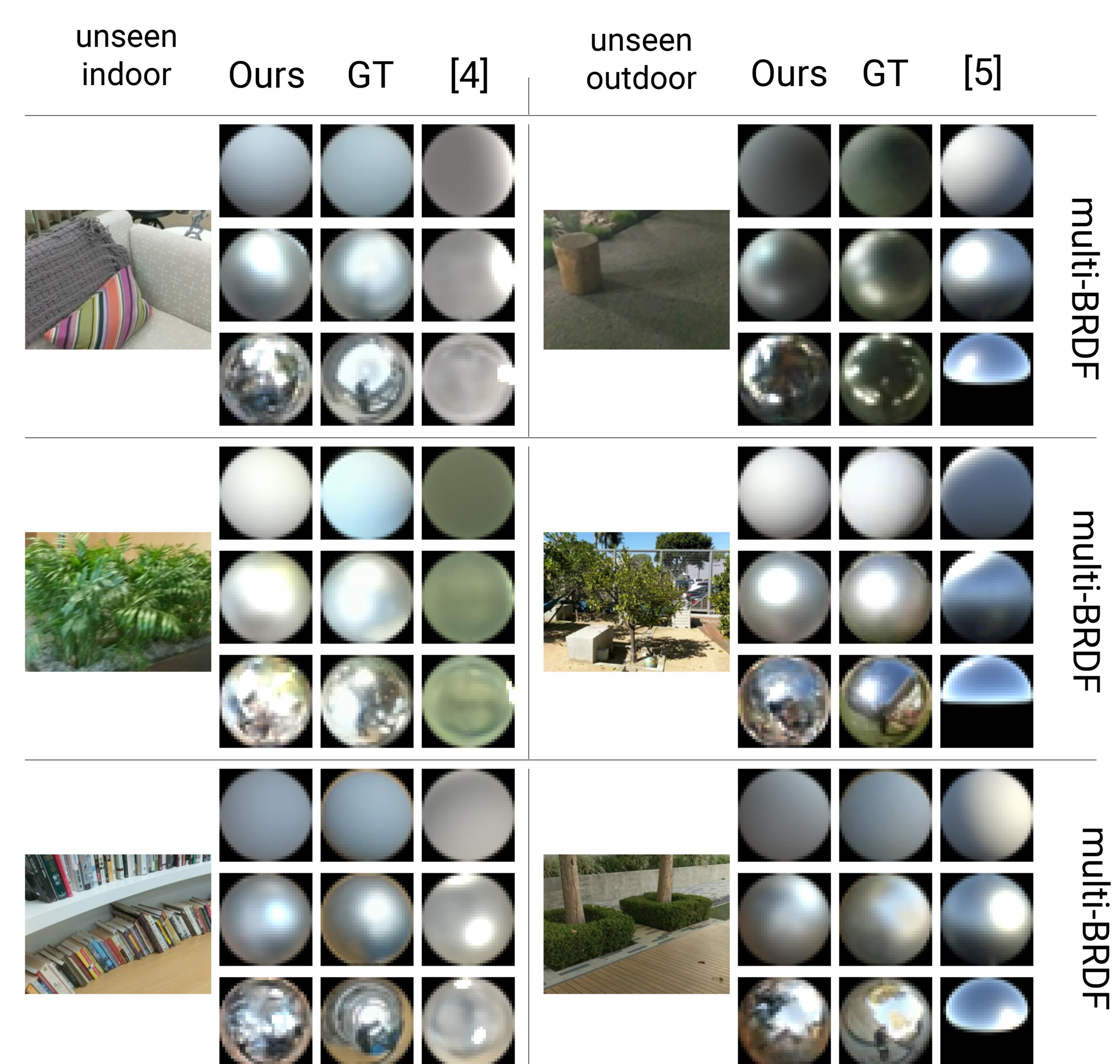
+ favor plausible lighting using LDR adversarial loss [3]

3. Results and comparisons

Renderings using our lighting inference (25^{th}/50^{th}/75^{th} percentile L_1 loss)

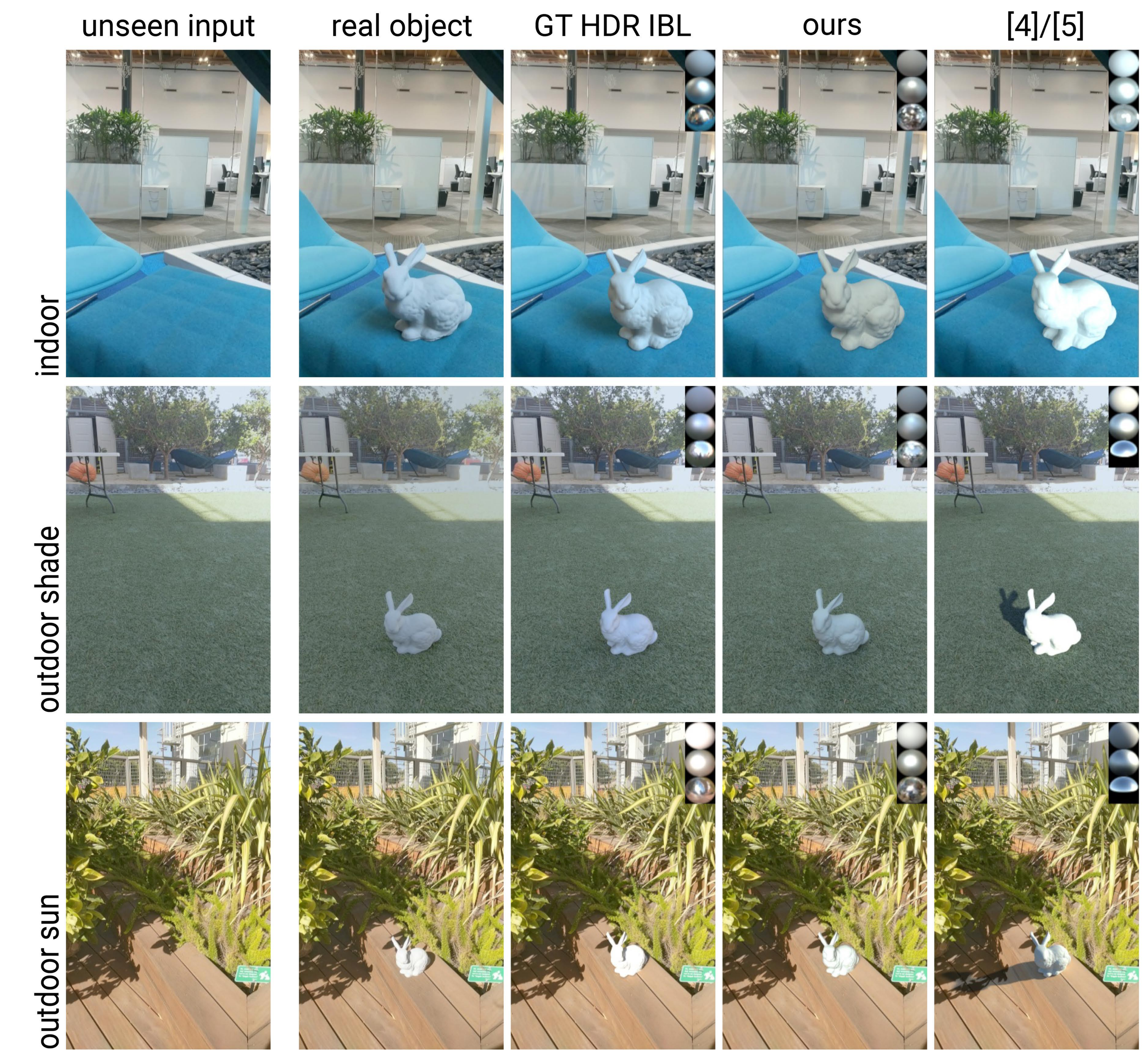


Comparisons with previous state-of-the-art [4] and [5]



Training on automatically exposed and white balanced images improves lighting estimation from a single image for AR, as compared to state-of-the-art for both indoor and outdoor scenes.

4. Comparing virtual object relighting



4. Runtime and mobile demo

Lightweight encoder using MobileNet v2 [6]
Lighting inference at **12-20 fps** on mobile CPU for real-time rendering

[Ask about our mobile demo!]

5. Summary

- An HDR lighting inference method for AR, trained using only LDR imagery, leveraging spheres with different materials to reveal different lighting cues in a single exposure.
- The first CNN-based lighting estimation approach to generalize to both indoor and outdoor scenes given a single input image.
- Improved lighting estimation for AR compared with previous work developed to handle only a single class of lighting.

References

[1] Debevec, Paul. "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography." ACM SIGGRAPH, 1998.

[2] Debevec, Paul, et al. "Acquiring the reflectance field of a human face." ACM SIGGRAPH, 2000.

[3] Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

[4] Gardner, Marc-André, et al. "Learning to Predict Indoor Illumination from a Single Image." ACM SIGGRAPH Asia, 2017.

[5] Hold-Geoffroy, Yannick et al. "Deep Outdoor Illumination Estimation." CVPR, 2017.

[6] Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." CVPR, 2018.